A renewal version of the Sanov theorem

Mauro Mariani, Lorenzo Zambotti

Electron. Commun. Probab. **19** (2014), no. 69, 1–13.
DOI: 10.1214/ECP.v19-3325
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

A renewal version of the Sanov theorem

Mauro Mariani * Lorenzo Zambotti†

Abstract

Large deviations for the local time of a process X_t are investigated, where $X_t = x_i$ for $t \in [S_{i-1}, S_i[$ and (x_j) are i.i.d. random variables on a Polish space, S_j is the j-th arrival time of a renewal process depending on (x_j). No moment conditions are assumed on the arrival times of the renewal process.

Keywords: Large deviations; Renewal processes, Sanov Theorem, Heavy tails.
AMS MSC 2010: 60K05; 60F10.
Submitted to ECP on February 16, 2014, final version accepted on October 8, 2014.

1 Main results

1.1 Outline of the result

Consider an i.i.d. sequence $(x_i)_{i\in\mathbb{N}^+}$ in a Polish space $\mathcal{X}$, with marginal distribution $\bar{\mu}$. One may define a stochastic process $(X_t)_{t\geq 0}$ on $\mathcal{X}$ by setting $X_t = x_i$ for $t \in [i-1, i[$, and consider its empirical measure $\pi_t := \frac{1}{t}\int_{[0,t[} ds\, \delta_{X_s}$. The ergodic theorem then states that $\pi_t \to \bar{\mu}$ as $t \to +\infty$, while the Sanov theorem yields a finer estimate for the probability that π_t is found in a small neighborhood of a given Borel probability measure $\bar{\nu}$ on $\mathcal{X}$. Such probability is estimated, in the sense of large deviations, as $\exp(-tH(\bar{\nu}|\bar{\mu}))$, where $H(\bar{\nu}|\bar{\mu})$ is the relative entropy of $\bar{\nu}$ with respect to $\bar{\mu}$.

In this paper, we want to provide a similar result, in the case in which the time spent by the process X_t at the point x_i may depend on the process itself. In particular, for $\tau\colon \mathcal{X} \to [0, +\infty]$ a measurable map, define $\mathcal{N}_t := \inf\{n \in \mathbb{N}^+ \;:\; \sum_{i=1}^{n+1} \tau(x_i) \geq t\}$, and $X_t := x_{\mathcal{N}_t+1}$. In the next section, the precise mathematical setting for the study of the large deviations of the empirical measure of X_t is recalled, and a large deviations result is established in Section 1.4. While for $\tau \equiv 1$ one gets the classical Sanov theorem, we are mainly interested in the case where the law of τ under $\bar{\mu}$ features heavy tails. In such a case the Markov process $(X_t, t - \sum_{i=1}^{\mathcal{N}_t} \tau(x_i))$ does not have good ergodic properties, and the classical Donsker-Varadhan theorem is violated.

1.2 Mathematical setting

In the following $\mathbb{N} = \{0, 1, \ldots\}$, $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$; $\mathcal{X}$ is a Polish space, that is a separable, completely metrisable topological space; a general element of $\mathcal{X}^{\mathbb{N}^+}$ will be denoted $\mathbf{x} = (x_1, x_2, \ldots)$; $C_b(\mathcal{X})$ and $C_c(\mathcal{X})$ are respectively the spaces of bounded continuous

*Dipartimento di Matematica, Università degli Studi di Roma La Sapienza, Italy.
E-mail: mariani@mat.uniroma1.t
†LPMA (CNRS UMR. 7599) Université Paris 6 – Pierre et Marie Curie, U.F.R. Mathématiques, France.
E-mail: lorenzo.zambotti@upmc.fr

functions and compactly supported continuous functions on $\mathcal{X}$. $\mathcal{M}_1(\mathcal{X})$ is the space positive Radon measure on $\mathcal{X}$ with total variation bounded by 1, while $\mathcal{P}(\mathcal{X}) \subset \mathcal{M}_1(\mathcal{X})$ is the set of Borel probability measures on $\mathcal{X}$. For $\mu \in \mathcal{M}_1(\mathcal{X})$ and f a μ-integrable function, we write $\mu(f) := \int d\mu\, f$. For $\mu,\, \nu \in \mathcal{P}(\mathcal{X})$, $\mathbf{H}(\nu|\mu)$ denotes the relative entropy of ν with respect to μ:

$$\mathbf{H}(\nu|\mu) := \sup_{\varphi \in C_b(\mathcal{X})} \nu(\varphi) - \log \mu(e^{\varphi}) = \begin{cases} \int \mu(dx) h\big(\frac{d\nu}{d\mu}\big) & \text{if } \nu << \mu; \\ +\infty & \text{otherwise}; \end{cases} \tag{1.1}$$

where the positive convex function h is defined as $h(\varrho) = \varrho\,(\log \varrho - 1) + 1$.

We always consider $\mathcal{P}(X)$ equipped with the narrow (or weak) topology, namely the weakest topology such that $\mu \mapsto \mu(f)$ is continuous for all $f \in C_b(\mathcal{X})$. In the particular case in which $\mathcal{X}$ is locally compact, we will also regard $\mathcal{M}_1(\mathcal{X})$ as a topological space, equipped with the vague topology, namely the weakest topology such that $\mu \mapsto \mu(f)$ is continuous for all $f \in C_c(\mathcal{X})$. $\mathcal{P}(X)$ is then a Polish space, and if $\mathcal{X}$ is locally compact $\mathcal{M}_1(\mathcal{X})$ is a compact Polish space.

Fix a reference probability $\bar{\mu} \in \mathcal{P}(\mathcal{X})$ and a measurable function $\tau\colon \mathcal{X} \to [0, +\infty]$; $\tau(x)$ has to be interpreted as the time elapsed at x. $\bar{\mu}$ and τ are the only 'inputs' of the problem.

Define $\xi\colon \mathcal{X} \to [0, +\infty]$ and $\xi^\infty \in [0, +\infty]$ as

$$\begin{aligned} \xi(x) &= \inf_{\delta>0}\, \sup \big\{ c \geq 0 \,:\, \bar{\mu}(e^{c\tau} \mathbb{1}_{B_\delta(x)}) < +\infty \big\} \\ \xi^\infty &:= \sup_{K \subset \mathcal{X},\, K \text{compact}} \sup \big\{ c \geq 0 \,:\, \bar{\mu}(e^{c\tau} \mathbb{1}_{K^c}) < +\infty \big\} \end{aligned} \tag{1.2}$$

where $B_\delta(x) \subset \mathcal{X}$ is the ball of radius δ centered at x, see (2.4) for another characterisation of ξ. Note $\xi^\infty = +\infty$ if $\mathcal{X}$ is compact.

The role of the auxiliary function ξ and of the assumptions below are discussed at the end of this section. In particular it is remarked that (A2) below is implied by regularity assumptions on τ (e.g. upper semicontinuity at infinity). Hereafter (A1) and (A2) will *always* be assumed, while our main results are proved whenever at least one of (A3) or (A4) holds (with somehow different statements in the two cases).

(A1) $\bar{\mu}(\{\tau = 0\}) = \bar{\mu}(\{\tau = +\infty\}) = 0$.

(A2) $\bar{\mu}(\{\xi < +\infty\}) = 0$.

(A3) $\xi^\infty = +\infty$.

(A4) $\mathcal{X}$ is locally compact.

In the following $\mathbf{x}$ is sampled as an i.i.d. sequence with marginal law $\bar{\mu}$ and $\mathbf{E}$ will denote the expectation of functions of $\mathbf{x}$ with respect to $\bar{\mu}^{\otimes \mathbb{N}^+}$. By (A1), for each $n \in \mathbb{N}$, $t \geq 0$ and a.e. $\mathbf{x}$, the following random variables are well defined

$$S_0 \equiv S_0(\mathbf{x}) := 0, \qquad S_n \equiv S_n(\mathbf{x}) := \sum_{i=1}^{n} \tau(x_i), \qquad n \geq 1,$$

$$\begin{aligned} \mathcal{N}_t \equiv \mathcal{N}_t(\mathbf{x}) &:= \inf\{n \in \mathbb{N} \,:\, S_{n+1} \geq t\} = \sum_{n=1}^{+\infty} \mathbb{1}_{(S_n \leq t)}, \\ X_t \equiv X_t(\mathbf{x}) &:= x_{\mathcal{N}_t+1}, \\ \pi_t \equiv \pi_t(\mathbf{x}) &= \frac{1}{t} \int_{[0,t[} ds\, \delta_{X_s} \in \mathcal{P}(\mathcal{X}). \end{aligned} \tag{1.3}$$

In other words, $X_t = x_1$ for $t \in [0, \tau(x_1)[$, $X_t = x_2$ for $t \in [\tau(x_1), \tau(x_1)+\tau(x_2)[$ and so on, while $\pi_t \colon \mathcal{X}^{\mathbb{N}^+} \to \mathcal{P}(\mathcal{X})$ is the local time or the empirical measure of X_t. Let $\mathbf{P}_t := \bar{\mu}^{\otimes \mathbb{N}^+} \circ \pi_t^{-1}$ be the law of π_t.

From the ergodic theorem, one expects π_t to concentrate on a deterministic limit as $t \to +\infty$ (this is easily established, for instance, whenever $\bar{\mu}(\tau) < +\infty$). Large deviations of $\mathbf{P}_t$ are then relevant, and subject of investigation of this paper.

1.3 Some examples

Taking advantage of the general metric setting, one is able to fit in this framework also the case of a process with random waiting time, see the examples (b) and (c) below.

(a) If $\tau(x) \equiv 1$, then we are in the framework of the classical Sanov theorem, [2, Chapter 6.2]. Here $\xi(x) = \xi^\infty = +\infty$ for all $x \in \mathcal{X}$.

(b) Assume $\mathcal{X} = \mathcal{Y} \times [0, +\infty]$ for some Polish space $\mathcal{Y}$. Let p be a Borel probability on $\mathcal{Y}$ and for p-a.e. y let ϕ_y be a probability on $[0, +\infty]$ concentrated on $]0, +\infty[$, with $y \mapsto \phi_y$ measurable. Set $d\bar{\mu}((y,t)) = dp(y)\, d\phi_y(t)$ and $\tau(y,t) = t$. Then we are in the framework of a pure jump process, jumping on $\mathcal{Y}$ with law p and spending a random time at a visited point y with law ϕ_y. In this case

$$\xi(y,t) = \begin{cases} \sup\{c \geq 0 \, : \int \phi_y(ds) e^{cs} < +\infty\} & \text{if } t = +\infty \text{ and } y \in \mathrm{Supp}(\nu) \\ +\infty & \text{otherwise.} \end{cases}$$

$$\xi^\infty = \sup_{K \subset \mathcal{Y},\, K \text{compact}} \inf_{y \in K^c} \xi(y, +\infty)$$

(c) As a special case of (b), take $\mathcal{X} := [0, +\infty[\times [0, +\infty]$ and for $\bar{\mu}(d(r,s)) = \nu(dr)\phi(ds)$, where ν is any probability measure on $]0, +\infty[$ and ϕ is the exponential law with mean 1. Set $\tau((y,s)) = \theta(y)s$, so that, conditionally on y, τ is an exponential random variable with mean $\theta(y)$. In this setting, $\mathcal{N}_t$ is an inhomogeneous Poisson random process, and the empirical measure π_t keeps track of the rates of the interarrival times. In this case $\xi(y,t) = +\infty$ for $t < +\infty$ or $y \notin \mathrm{Supp}(\nu)$, while $\xi(y, +\infty) = 1/\theta(y)$ for $y \in \mathrm{Supp}(\nu)$, and $\xi^\infty = \underline{\lim}_{y \to +\infty} \xi(y, +\infty)$.

(d) An interesting example in which τ is 'truly' deterministic is the following. $\mathcal{X} =]0, +\infty[^n$, $\bar{\mu}(dx) = \prod_{i=1}^n \bar{\mu}_i(dx_i)$ for some probabilities $\bar{\mu}_i \in \mathcal{P}([0, +\infty[)$ and $\tau(x) = \frac{1}{n}\sum_{i=1}^n \frac{1}{x_i}$. This is a model for a particle moving on 1-dimensional torus of length 1. During its motion the particle touches some fixed *hot* points equi-spaced on the torus, and it changes its speed by sampling a new one with law $\bar{\mu}_i$ at the hot point i. $\tau(x)$ is then the time elapsed to complete a tour of the torus.

One can derive the large deviations of some physical quantities (e.g. kinetic energy of the particle) from the large deviations of the empirical measure of X_t. The physically relevant case is $\bar{\mu}_i(x_i) = x_i e^{-\beta_i x_i^2} dx_i$ for some $\beta_i > 0$. Then $\xi^\infty = +\infty$ and $\xi(x) = +\infty$ unless one the x_i is 0, in which case $\xi(x) = 0$. As remarked below, when $\{\xi = 0\}$ is non-empty, the large deviations rate functional is not strictly convex. For $n = 1$, this moving particle dynamics has been used as a building block of a toy model of out-of-equilibrium statistical mechanics in [6], where the absence of strict convexity of the rate causes a dynamic phase transition in the model.

1.4 Large deviations results

We recall the following standard definition.

Definition 1.1. *Let $\mathcal{Y}$ be Polish space and $(\mathbf{Q}_t)_{t>0}$ a family of Borel probability measures on $\mathcal{Y}$ and $I \colon \mathcal{Y} \to [0, +\infty]$. Then:*

- *I is* good *if $\{y \in \mathcal{Y} \,:\, I(y) \le M\}$ is compact in $\mathcal{Y}$ for all $M > 0$ and $I \not\equiv +\infty$.*
- *$(\mathbf{Q}_t)_{t>0}$ satisfies a* large deviations upper bound *with good rate I if*

$$\overline{\lim_{t\to+\infty}} \frac{1}{t} \log \mathbf{Q}_t(\mathcal{C}) \le - \inf_{u\in\mathcal{C}} I(u) \qquad \textit{for all } \mathcal{C} \subset \mathcal{Y} \textit{ closed.}$$

- *$(\mathbf{Q}_t)_{t>0}$ satisfies a* large deviations lower bound *with good rate I, if*

$$\underline{\lim_{t\to+\infty}} \frac{1}{t} \log \mathbf{Q}_t(\mathcal{O}) \ge - \inf_{u\in\mathcal{O}} I(u) \qquad \textit{for all } \mathcal{O} \subset \mathcal{Y} \textit{ open.}$$

$(\mathbf{P}_t)_{t>0}$ is said to satisfy a good large deviations principle *if both the upper and lower bounds hold with the same good rate I.*

For $\nu \in \mathcal{M}_1(\mathcal{X})$, let ν_a and ν_s be respectively the absolutely continuous and singular parts of ν with respect to $\bar{\mu}$. If ν is such that $\nu(1/\tau) \in]0, +\infty[$ define $\bar{\nu} \in \mathcal{P}(\mathcal{X})$ as

$$\bar{\nu}(dx) = \frac{\frac{1}{\tau(x)}\nu(dx)}{\nu(1/\tau)}. \tag{1.4}$$

Proposition 1.2. *Define $I\colon \mathcal{P}(\mathcal{X}) \to [0, +\infty]$ as*

$$I(\nu) = \begin{cases} \nu_a(1/\tau)\mathbf{H}(\bar{\nu}_a|\bar{\mu}) + \nu_s(\xi) & \textit{if } \nu_a(1/\tau) < +\infty, \\ +\infty & \textit{otherwise,} \end{cases}$$

where we define $\nu_a(1/\tau)\mathbf{H}(\bar{\nu}_a|\bar{\mu}) = 0$ whenever $\nu_a(1/\tau) = 0$. If (A3) holds, then I is a good and convex functional on $\mathcal{P}(\mathcal{X})$.

Theorem 1.3. *If (A3) holds, then $(\mathbf{P}_t)_{t>0}$ satisfies a good large deviations principle on $\mathcal{P}(\mathcal{X})$ with rate I.*

In the following remark some features of the functional I are investigated. In particular we characterise the cases where I is strictly convex and those in which it features affine stretches.

Remark 1.4. Assume (A3). Since $\xi(x) = +\infty$ for $x \notin \mathrm{Supp}(\bar{\mu})$, $I(\nu) = +\infty$ if $\mathrm{Supp}(\nu) \not\subset \mathrm{Supp}(\bar{\mu})$. However, contrary to classical Sanov theorem, in general $I(\nu) < +\infty$ does not imply that ν is absolutely continuous with respect to $\bar{\mu}$, unless $\xi \equiv \infty$. In general, the nature of $I(\nu)$ depends on the values of ξ and $\bar{\mu}(\tau)$. Indeed let

$$E := \{x \in \mathcal{X} \,:\, \xi(x) = 0\}$$

be the set of points around which τ has no local exponential moments. Then

(1) If $E = \emptyset$, namely if $\xi(x) > 0$ for all $x \in \mathcal{X}$, then a fortiori $\bar{\mu}(\tau) < +\infty$ and $I(\nu) = 0$ iff $\nu = \mu$, where (consistently with (1.4))

$$\mu(dx) := \frac{\tau(x)\bar{\mu}(dx)}{\bar{\mu}(\tau)}. \tag{1.5}$$

(2) If $E \neq \emptyset$, there are two possibilities

  (2A) If $\bar{\mu}(\tau) < +\infty$, then $I(\nu) = 0$ iff $\nu = \alpha\mu + (1-\alpha)\lambda$ for some $\alpha \in [0, 1]$ and some $\lambda \in \mathcal{P}(X)$ such that $\lambda(E) = 1$, where μ is given by (1.5).
  (2B) If $\bar{\mu}(\tau) = +\infty$ then $I(\nu) = 0$ iff ν is concentrated on E.

In particular, Theorem 1.3 implies the convergence in law of π_t to μ in case (1), and in case (2B) if E is a singleton. In all other cases, a nontrivial second order large deviations may hold, see [10] where moderate deviations are discussed in a particular case. Finally, if $E \neq \emptyset$, then the subdifferential of I is nontrivial.

If $\xi^\infty < +\infty$, $(\mathbf{P}_t)_{t>0}$ is not exponentially tight on $\mathcal{P}(\mathcal{X})$, and large deviations need to be investigated on $\mathcal{M}_1(\mathcal{X})$. However, in this case we need $\mathcal{X}$ to be locally compact in order to have good topological properties of $\mathcal{M}_1(\mathcal{X})$.

Proposition 1.5. *Define $I' \colon \mathcal{M}_1(\mathcal{X}) \to [0, +\infty]$ as*

$$I'(\nu) = \begin{cases} \nu_a(1/\tau)\mathbf{H}(\bar{\nu}_a|\bar{\mu}) + \nu_s(\xi) + (1-\nu(\mathcal{X}))\xi^\infty & \text{if } \nu_a(1/\tau) < +\infty, \\ +\infty & \text{otherwise,} \end{cases}$$

If (A4) holds, then I' is a good and convex functional on $\mathcal{M}_1(\mathcal{X})$.

Theorem 1.6. *If (A4) holds, then $(\mathbf{P}_t)_{t>0}$ satisfies a good large deviations principle on $\mathcal{M}_1(\mathcal{X})$ with rate I'.*

Under (A1), the key assumption (A2) is satisfied whenever

$$\bar{\mu}\big(\cap_{M>0} \operatorname{Closure}(\{\tau \geq M\})\big) = 0.$$

In particular (A2) holds if τ is upper semicontinuous at infinity. Since all the results stated above make sense even dropping (A2), one may wonder whether it is a merely technical condition. While one can prove the large deviations upper bound even dropping this assumption, the lower bound is in general false if (A2) does not hold.

1.5 Outlook

With the same notation as above, one may also introduce the Markov process $Y_t = (X_t, \frac{t - \sum_{i=1}^{\mathcal{N}_t} \tau(x_i)}{\tau(x_{\mathcal{N}_t+1})}) \in \mathcal{X} \times [0, 1[$. Large deviations for the empirical measure of Y_t would give large deviations of X_t by a standard contraction argument. Moreover, the Donsker-Varadhan theory [3] and its extensions provide general large deviations results for the empirical measure of a Markov process. However, this approach fails in this case. On the one hand, standard Donsker-Varadhan theorems cannot be applied here, since Y_t only enjoys weak ergodic properties. On the other hand, even formally, the Donsker-Varadhan rate functional does not provide the right answer, a feature already remarked in [5] for renewal processes. Indeed, it has been proved in [7] that in general the empirical measure of Y_t *does not satisfy* a large deviations principle, and in the special case it does (which depends on the law of τ under $\bar{\mu}$), the rate functional does not correspond to the Donsker-Varadhan functional. Similarly, the large deviations rate functional for π_t does not correspond in general to the one predicted by applying contraction to the Donsker-Varadhan functional for the empirical measure of Y_t (unless τ has all exponential moments bounded). In this respect, it may be remarkable that the law of π_t satisfies a large deviations principle at all.

2 The functional I

This section is devoted to prove Proposition 1.2, Proposition 1.5 and general properties of the functional I, which will play a key role in the proof of the main theorems. First we remark that one can reduce to the case of a compact state space $\mathcal{X}$.

Proposition 2.1. *Suppose that Proposition 1.2 and Theorem 1.3 hold with the additional hypotheses of $\mathcal{X}$ being a* compact *Polish space. Then Proposition 1.2, Theorem 1.3, Proposition 1.5 and Theorem 1.6 hold.*

Proof. An arbitrary Polish space $\mathcal{X}$ embeds continuously in the compact Polish space $[0,1]^{\mathbb{N}}$, see [9, Lemma 3.1.2]. Regard $\mathcal{X}$ as a subset of $[0,1]^{\mathbb{N}}$ and let $\mathcal{Y}$ be the closure of $\mathcal{X}$. Then $\mathcal{Y}$ is compact. Extend $\bar{\mu}$ to $\mathcal{Y}$ setting $\bar{\mu}(\mathcal{Y} \setminus \mathcal{X}) = 0$ and extend τ to $\mathcal{Y}$ setting $\tau(x) = +\infty$ for $x \in \mathcal{Y} \setminus \mathcal{X}$. We denote $\xi_{\mathcal{Y}}$ and $I_{\mathcal{Y}}$ the object corresponding to ξ and I on

$\mathcal{Y}$. Then (A1), (A2) hold on $\mathcal{Y}$ since they hold on $\mathcal{X}$, while refa3 is trivially satisfied on $\mathcal{Y}$. Thus, by the hypotheses of this proposition, the extension of $\mathbf{P}_t$ to $\mathcal{P}(\mathcal{Y})$ satisfies a large deviations principle with good rate $I_{\mathcal{Y}}$. We then separate the two cases, wether (A3) or (A4) hold on $\mathcal{X}$.

If (A3) holds (on $\mathcal{X}$), then $\xi_{\mathcal{Y}}(x) = +\infty$ for $x \in \mathcal{Y} \setminus X$ (since neighborhoods of such points x in $\mathcal{Y}$ are exactly complements of compact subsets of $\mathcal{X}$). Thus the map $\Pi \colon \mathcal{P}(\mathcal{Y}) \to \mathcal{P}(\mathcal{X})$ defined as

$$\Pi(\nu) = \begin{cases} \nu(\cdot|\mathcal{X}) := \frac{\nu(\cdot \cap \mathcal{X})}{\nu(\mathcal{X})} & \text{if } \nu(\mathcal{X}) > 0 \\ \bar{\mu} & \text{otherwise} \end{cases}$$

is continuous on the domain of $I_{\mathcal{Y}}$. Since Π is just the restriction map for probabilities concentrated on $\mathcal{X}$, the extension of $\mathbf{P}_t$ to $\mathcal{P}(\mathcal{Y})$ is mapped to $\mathbf{P}_t$ by Π. Then by contraction principle [2, Chapter 4.2], I is good and $\mathbf{P}_t$ satisfies a good large deviations principle on $\mathcal{P}(\mathcal{X})$ with rate I. It is immediate to check that Π preserves the convexity, so I is convex.

Suppose now (A4) holds (but not (A3)). Consider the map $\Pi' \colon \mathcal{P}(\mathcal{Y}) \to \mathcal{M}_1(\mathcal{X})$ defined by

$$\Pi'(\nu)(f) = \nu(f) \qquad \forall f \in C_c(\mathbf{X})$$

where we also identified f with its unique continuous extension on $\mathcal{Y}$ (namely $f(x) = 0$ for $x \in \mathcal{Y} \setminus \mathcal{X}$). Then Π' is continuous, and we conclude again by contraction principle. □

Motivated by the previous remark, hereafter we assume $\mathcal{X}$ to be compact, with no loss of generality.

For $\delta > 0$, define $\xi_\delta \colon \mathcal{X} \to [0, +\infty]$ as

$$\xi_\delta(x) = \sup\left\{c \,:\, \bar{\mu}(e^{c\tau} \mathbb{1}_{B_\delta(x)}) < +\infty\right\} \tag{2.1}$$

In particular $\xi = \sup_{\delta>0} \xi_\delta$. Let $\hat{\xi}_\delta$ be the lower semicontinuous envelope of ξ_δ.

Lemma 2.2. *For all $x \in \mathcal{X}$, $\xi(x) = \sup_{\delta>0} \hat{\xi}_\delta(x)$. In particular ξ is lower semicontinuous.*

Proof. By the very definition of ξ_δ, if $y \in B_\delta(x)$, then $\xi_{2\delta}(x) \le \xi_\delta(y)$. Therefore

$$\xi_\delta(x) \ge \hat{\xi}_\delta(x) := \sup_{\varepsilon>0} \inf_{y \in B_\varepsilon(x)} \xi_\delta(y) \ge \inf_{y \in B_\delta(x)} \xi_\delta(y) \ge \xi_{2\delta}(x)$$

The lemma follows taking the supremum in $\delta > 0$. □

Let $LSC(\mathcal{X})$ be the set of lower semicontinuous functions $f : \mathcal{X} \to]-\infty, +\infty]$. If $f \in LSC(\mathcal{X})$ then f is bounded from below.

Lemma 2.3. *Recall* (2.1). *For all $M < +\infty$ and ε, $\delta > 0$ (hereafter $a \wedge b := \min(a, b)$)*

$$\bar{\mu}(e^{(\xi_\delta \wedge M - \varepsilon)\tau}) < +\infty. \tag{2.2}$$

On the other hand, if $f \in LSC(\mathcal{X})$ is such that

$$\bar{\mu}(e^{\tau f}) < +\infty, \tag{2.3}$$

then $f(x) \le \xi(x)$ for all $x \in \mathcal{X}$. *In particular*

$$\xi(x) = \sup\{f(x),\ f \in LSC(\mathcal{X}) \,:\, \bar{\mu}(e^{\tau f}) < +\infty\}. \tag{2.4}$$

Proof. Fix M, ε, $\delta > 0$ and let $\{B_{\delta/2}(y_1), \dots, B_{\delta/2}(y_n)\}$ be a finite covering of the compact space $\mathcal{X}$ with balls of radius $\delta/2$. Since $\xi_\delta(x) \le \xi_{\delta/2}(y_i)$ for $x \in B_{\delta/2}(y_i)$

$$\bar\mu(e^{(\xi_\delta \wedge M - \varepsilon)\tau}) \le \sum_{i=1}^n \bar\mu(e^{(\xi_\delta \wedge M-\varepsilon)\tau} \mathbb{1}_{B_{\delta/2}(y_i)}) \le \sum_{i=1}^n \bar\mu(e^{(\xi_{\delta/2}(y_i)\wedge M-\varepsilon)\tau} \mathbb{1}_{B_{\delta/2}(y_i)}).$$

Since $\xi_{\delta/2}(y_i) \wedge M - \varepsilon < \xi_{\delta/2}(y_i)$, each term in the summation in the last line of the above formula is finite by the very definition of $\xi_{\delta/2}(y_i)$. Thus (2.2) holds.

Let now $f \in LSC(\mathcal{X})$, and suppose that for some $x \in \mathcal{X}$ and $\varepsilon > 0$, $f(x) \ge \xi(x) + 2\varepsilon$. Since f is lower semicontinuous, there exists $\delta > 0$ such that $\inf_{y \in B_\delta(x)} f(y) \ge \xi(x) + \varepsilon$. Then

$$\bar\mu(e^{\tau f}) \ge \bar\mu(e^{\tau f}\mathbb{1}_{B_\delta(x)}) \ge \bar\mu(e^{\tau[\xi(x)+\varepsilon]}\mathbb{1}_{B_\delta(x)}) \ge \bar\mu(e^{\tau[\xi_\delta(x)+\varepsilon]}\mathbb{1}_{B_\delta(x)}) = +\infty.$$

Therefore if (2.3) holds, then $f \le \xi$ everywhere. □

Proposition 2.4. *For each* $\nu \in \mathcal{P}(\mathcal{X})$

$$I(\nu) = \sup\big\{\nu(f),\ f \in LSC(\mathcal{X})\ :\ \bar\mu(e^{\tau f}) \le 1\big\} =: \tilde I(\nu). \tag{2.5}$$

In particular Proposition 1.2 holds.

Proof. Fix $\nu \in \mathcal{P}(\mathcal{X})$, and let $f\colon \mathcal{X} \to \mathbb{R}$ be Borel measurable, ν-integrable, such that $\bar\mu(e^{\tau f}) < 1$ and $f \le (\hat\xi_\delta \wedge M - \varepsilon)$ for some M, δ, $\varepsilon > 0$. Since continuous functions are dense in $L_1(\nu + \bar\mu)$, there exists a sequence (f_n) in $LSC(\mathcal{X})$ such that $f_n \to f$ in $L_1(d\nu)$ and (up to passing to a subsequence) also $\bar\mu$-almost everywhere. Moreover one can assume $f_n \le \hat\xi_\delta \wedge M - \varepsilon$, since the sequence $f_n \wedge (\hat\xi_\delta \wedge M - \varepsilon)$ is in $LSC(\mathcal{X})$ and enjoys the aforementioned properties as well. Dominated convergence and (2.2) imply $\lim_n \bar\mu(e^{\tau f_n}) = \bar\mu(e^{\tau f}) < 1$. Therefore $\bar\mu(e^{\tau f_n}) \le 1$ for n large enough. Thus

$$\tilde I(\nu) \ge \sup_{M,\,\delta,\,\varepsilon>0} \sup\big\{\nu(f),\ f\ \nu\text{-integrable such that}\quad \bar\mu(e^{\tau f}) < 1,\ f \le \hat\xi_\delta \wedge M - \varepsilon\big\}. \tag{2.6}$$

By (A2), the Borel set $A = \{\xi = +\infty\}\setminus \mathrm{Supp}(\nu_s)$ is such that $\bar\mu$ and ν_a are concentrated on A and ν_s is concentrated on A^c. Fix M, δ, $\varepsilon > 0$ and take $\varphi \in C(\mathcal{X})$ such that $\bar\mu(e^\varphi) \le 1$. In the right hand side of (2.6) consider a f of the form

$$f = \big(\frac{\varphi}{\tau} \wedge \hat\xi_\delta \wedge M\big)\mathbb{1}_A + (\hat\xi_\delta \wedge M)\mathbb{1}_{A^c} - \varepsilon. \tag{2.7}$$

Then $\bar\mu(e^{\tau f}) = \bar\mu(e^{\tau f}\mathbb{1}_A) \le \bar\mu(e^{\varphi-\varepsilon}) \le e^{-\varepsilon} < 1$.

If $\nu_a(1/\tau) = +\infty$, take $\varphi \equiv 1$ in (2.7). Then f is ν-integrable and by monotone convergence $\nu(f) \to +\infty$ as one lets $M \to +\infty$ and $\delta \downarrow 0$, so that $\tilde I(\nu) = +\infty$ by (2.6). Thus $\tilde I(\nu) = I(\nu) = +\infty$ whenever $\nu_a(1/\tau) = +\infty$.

Consider then the case $\nu_a(1/\tau) < +\infty$. Since φ is bounded, any f of the form (2.7) is ν-integrable, and thus by (2.6)

$$\tilde I(\nu) \ge \nu(f) = \nu_a\big(\frac{\varphi}{\tau} \wedge \hat\xi_\delta \wedge M\big) + \nu_s(\hat\xi_\delta \wedge M) - \varepsilon.$$

Take the limit $M \to +\infty$, $\delta \downarrow 0$, $\varepsilon \downarrow 0$. Monotone convergence and Lemma 2.2 then yield

$$\tilde I(\nu) \ge \nu_a\big(\frac{\varphi}{\tau}\big) + \nu_s(\sup_{\delta>0}\hat\xi_\delta) = \nu_a\big(\frac{\varphi}{\tau}\big) + \nu_s(\xi) = \nu_a(1/\tau)\bar\nu_a(\varphi) + \nu_s(\xi)$$

where the last equality is a direct consequence of (1.4). Now optimize over φ to get

$$\begin{aligned}\tilde I(\nu) &\ge \nu_a(1/\tau)\sup\big\{\bar\nu_a(\varphi),\ \varphi \in C(\mathcal{X})\ :\ \bar\mu(e^\varphi) \le 1\big\} + \nu_s(\xi)\\ &\ge \nu_a(1/\tau)\sup\big\{\bar\nu_a(\varphi) - \log\bar\mu(e^\varphi),\ \varphi \in C(\mathcal{X})\ :\ \bar\mu(e^\varphi) = 1\big\} + \nu_s(\xi)\end{aligned}$$

Notice that the condition $\bar\mu(e^\varphi)=1$ can now be dropped in the supremum in the last line above, since for any $c\in\mathbb{R}$ the change $\varphi\mapsto\varphi+c$ leaves the quantity $\bar\nu_a(\varphi)-\log\bar\mu(e^\varphi)$ invariant. Therefore the supremum over φ equals the relative entropy as defined in (1.1), so that $\tilde I\geq I$.

In order to prove $I(\nu)\geq\tilde I(\nu)$, one only needs to consider the case $\nu_a(1/\tau)<+\infty$, the inequality being trivial otherwise. Then for $\varphi\in L_1(d\bar\nu_a)$ such that $\bar\mu(e^\varphi)\leq 1$,

$$\nu_a(1/\tau)\mathbf{H}(\bar\nu_a|\bar\mu)\geq\nu_a(1/\tau)\big[\bar\nu_a(\varphi)-\log\bar\mu(e^\varphi)\big]\geq\nu_a(\varphi/\tau)=\nu_a(f),$$

where $f:=\varphi/\tau$ and the above conditions on φ translates into $f\in L_1(d\nu_a)$ and $\bar\mu(e^{\tau f})\leq 1$. Therefore, optimizing over $f\in LSC(\mathcal{X})$ satisfying these two conditions, and noting that Lemma 2.3 implies $f\leq\xi$ for such a f

$$\begin{aligned}I(\nu)&=\nu_a(1/\tau)\mathbf{H}(\bar\nu_a|\bar\mu)+\nu_s(\xi)\\&\geq\sup\{\nu_a(f),\,f\in LSC(\mathcal{X})\cap L_1(d\nu_a)\,:\,\bar\mu(e^{\tau\,f})\leq 1\}+\nu_s(\xi)\\&=\sup\{\nu_a(f)+\nu_s(\xi),\,f\in LSC(\mathcal{X})\,:\,\bar\mu(e^{\tau\,f})\leq 1\}\\&\geq\sup\{\nu_a(f)+\nu_s(f),\,f\in LSC(\mathcal{X})\,:\,\bar\mu(e^{\tau\,f})\leq 1\}=\tilde I(\nu).\end{aligned}$$

Now (2.5) states in particular that I is the supremum of a family of linear lower semicontinuous mappings, thus Proposition 1.2 follows. □

Lemma 2.5. *For $A\subset\mathcal{X}$ a Borel set, define*

$$\xi^A:=\sup\big\{c\geq 0\,:\,\bar\mu(e^{c\tau}\mathbb{1}_A)<+\infty\big\},$$

$$\underline{\xi}^A:=-\varlimsup_{L\to+\infty}\frac{1}{L}\log\bar\mu\big(\{\tau\geq L\}\cap A\big).$$

Then $\underline{\xi}^A=\xi^A$.

Proof. For $c>0$

$$\bar\mu(e^{c\tau}\mathbb{1}_A)=\int_{\mathbb{R}^+}d\eta\,\bar\mu(\{e^{c\tau}\geq\eta\}\cap A)=c\int_{\mathbb{R}^+}dL\,\bar\mu(\{\tau\geq L\}\cap A)\,e^{c\,L}.$$

It is then easy to check that, for $c>\underline{\xi}^A$, $\bar\mu(e^{c\tau}\mathbb{1}_A)=+\infty$, while if $\xi^A>0$ and $0<c<\xi^A$, then $\bar\mu(e^{c\tau}\mathbb{1}_A)<+\infty$. It follows $\xi^A=\underline{\xi}^A$. □

Proposition 2.6. *Define $J:\mathcal{P}(\mathcal{X})\to[0,+\infty]$ as*

$$J(\nu)=\begin{cases}I(\nu)&\text{if }\nu=\nu_a,\\+\infty&\text{otherwise.}\end{cases}\tag{2.8}$$

I is the lower semicontinuous envelope of J.

Notice that in the classical case $\tau\equiv 1$, J coincides with I. However, in this general case, $I=J$ iff $\xi\equiv+\infty$.

Proof of Proposition 2.6. Since $J\geq I$ and I is lower semicontinuous, the lower semicontinuous envelope of J is greater than I. Therefore it is enough to show that for each $\nu\in\mathcal{P}(\mathcal{X})$ such that $I(\nu)<+\infty$, there exists a sequence $\nu^n\to\nu$ such that $\varlimsup_n J(\nu^n)\leq I(\nu)$.

Let $\nu=\nu_a+\nu_s$ satisfy $I(\nu)<+\infty$. Since $\mathcal{X}$ is compact, for each $\delta\in(0,1)$ there exist $n^\delta\in\mathbb{N}^+$ and a finite Borel partition $(A_1^\delta,\dots,A_{n^\delta}^\delta)$ of $\mathcal{X}$ such that each A_i^δ has diameter bounded by δ, has nonempty interior, and satisfies $\nu_s(\partial A_i^\delta) = 0$. For $\delta > 0$ and $M > L \geq 0$, define

$$A_i^{\delta,L,M} := \{L \leq \tau \leq M\} \cap A_i^\delta.$$

Fix a $j \in \{1, \dots, n^\delta\}$. We claim that

$$\text{if } \nu_s(A_j^\delta) > 0 \text{ then } \forall L \geq 0,\ \exists M^L \geq L \text{ such that } \bar{\mu}(A_j^{\delta,L,M}) > 0 \text{ for all } M \geq M^L. \tag{2.9}$$

Indeed $\nu_s(\xi) \leq I(\nu) < +\infty$, thus ν_s is concentrated on $\{\xi < +\infty\}$. Since $\nu_s(\partial A_j^\delta) = 0$, there exists a point x_j^δ in the interior of A_j^δ such that $\xi(x_j^\delta) < +\infty$. Then, for each $c > \xi(x_j^\delta)$ and $\varepsilon > 0$

$$\lim_{M\to+\infty} \bar{\mu}\big(e^{c\tau}\mathbb{1}_{B_\varepsilon(x_j^\delta)}\mathbb{1}_{L\leq\tau\leq M}\big) = \bar{\mu}(e^{c\tau}\mathbb{1}_{B_\varepsilon(x_j^\delta)}\mathbb{1}_{\tau\geq L}) = +\infty.$$

Hence for M large enough $\{L \leq \tau \leq M\}$ has positive $\bar{\mu}$-measure in each neighbourhood of x_j^δ, including A_j^δ. The claim (2.9) is thus proved.

By (2.9), for each $\mathbf{L} = (L_1, L_2, \dots) \in [0, +\infty[^{\mathbb{N}}$ there exists $\mathbf{M^L} \in [0, +\infty[^{\mathbb{N}}$, such that the probability measure

$$\nu^{\delta,\mathbf{L},\mathbf{M}}(dx) := \nu_a(dx) + \sum_{i=1}^{n^\delta} \nu_s(A_i^\delta) \frac{\tau(x)\bar{\mu}(dx|A_i^{\delta,L_i,M_i})}{\bar{\mu}(\tau|A_i^{\delta,L_i,M_i})} \tag{2.10}$$

is well defined whenever $\mathbf{M} \geq \mathbf{M^L}$, provided the terms in the summation are understood to vanish whenever $\nu_s(A_i^\delta)$ does. It follows straightforwardly from this definition that for each $\varphi \in C_b(\mathcal{X})$

$$\begin{aligned}
&\lim_{\delta\downarrow 0} \sup_{\mathbf{L}\in[0,+\infty[^{\mathbb{N}},\, \mathbf{M}\geq\mathbf{M^L}} \left|\nu^{\delta,\mathbf{L},\mathbf{M}}(\varphi) - \nu(\varphi)\right| \\
&\quad\leq \varlimsup_{\delta\downarrow 0} \sum_{i=1}^{n^\delta} \nu_s(A_i^\delta)\big[\sup_{x\in A_i^\delta}\varphi(x) - \inf_{x\in A_i^\delta}\varphi(x)\big] = 0.
\end{aligned} \tag{2.11}$$

Note that for each $\delta > 0$ and $\mathbf{L}, \mathbf{M} \in [0, +\infty[^{\mathbb{N}}$ with $\mathbf{M} \geq \mathbf{M}^L$, $\nu^{\delta,\mathbf{L},\mathbf{M}}$ is absolutely continuous with respect to $\bar{\mu}$. By the convexity of I proved in Proposition 2.4

$$\begin{aligned}
J(\nu^{\delta,\mathbf{L},\mathbf{M}}) =& I(\nu^{\delta,\mathbf{L},\mathbf{M}}) \leq \nu_a(\mathcal{X}) I\Big(\frac{1}{\nu_a(\mathcal{X})}\nu_a\Big) \\
&\qquad + \sum_{i=1}^{n^\delta} \nu_s(A_i^\delta) I\Big(\frac{\tau(x)\bar{\mu}(dx|A_i^{\delta,L_i,M_i})}{\bar{\mu}(\tau|A_i^{\delta,L_i,M_i})}\Big) \\
&= I(\nu) - \Big[\nu_s(\xi) - \sum_{i=1}^{n^\delta} \nu_s(A_i^\delta) I\Big(\frac{\tau(x)\bar{\mu}(dx|A_i^{\delta,L_i,M_i})}{\bar{\mu}(\tau|A_i^{\delta,L_i,M_i})}\Big)\Big]
\end{aligned} \tag{2.12}$$

where the corresponding terms above are understood to vanish whenever $\nu_a(\mathcal{X})$ or $\nu_s(A_i^\delta)$ do. By direct computation

$$\begin{aligned}
I\Big(\frac{\tau(x)\bar{\mu}(dx|A_i^{\delta,L_i,M_i})}{\bar{\mu}(\tau|A_i^{\delta,L_i,M_i})}\Big) &= -\frac{1}{\bar{\mu}(\tau|A_i^{\delta,L_i,M_i})} \log \bar{\mu}(A_i^{\delta,L_i,M_i}) \\
&\leq -\frac{1}{L_i} \log \bar{\mu}(\{L_i \leq \tau \leq M_i\} \cap A_i^\delta).
\end{aligned}$$

Thus, from Lemma 2.5

$$\varlimsup_{L_i\to+\infty} \varlimsup_{M_i\to+\infty} \Big(\frac{\tau(x)\bar{\mu}(dx|A_i^{\delta,L_i,M_i})}{\bar{\mu}(\tau|A_i^{\delta,L_i,M_i})}\Big) \leq \xi^{A_i^\delta}.$$

Now, since $\xi \geq \xi^{A_i^\delta}$ on A_i^δ

$$\varlimsup_{\mathbf{L}\to+\infty} \varlimsup_{\mathbf{M}\to+\infty} \sum_{i=1}^{n^\delta} \nu_s(A_i^\delta) I\Big(\frac{\tau(x)\bar{\mu}(dx|A_i^{\delta,L_{i,k}^\delta})}{\bar{\mu}(\tau|A_i^{\delta,L_{i,k}^\delta})}\Big)\Big] \leq \sum_{i=1}^{n^\delta} \nu_s(A_i^\delta)\xi^{A_i^\delta} \leq \nu_s(\xi).$$

Together with (2.12) this implies

$$\sup_{\delta>0} \varlimsup_{\mathbf{L}\to+\infty} \varlimsup_{\mathbf{M}\to+\infty} J(\nu^{\delta,\mathbf{L},\mathbf{M}}) \leq I(\nu).$$

Combining this with (2.11), by a standard diagonal argument, there exists a sequence $\nu^n = \nu^{\delta^n,\mathbf{L}^n,\mathbf{M}^n}$ converging to ν such that $\varlimsup_n I(\nu^n) \leq I(\nu)$. □

3 Large deviations of the empirical measure

The following identity follows immediately from (1.3), and will come handy in this section.

$$\pi_t = \frac{1}{t}\sum_{i=1}^{\mathcal{N}_t} \tau(x_i)\delta_{x_i} + \frac{t - S_{\mathcal{N}_t}}{t}\delta_{x_{\mathcal{N}_t}+1}. \tag{3.1}$$

Lemma 3.1. *Let $f\colon \mathcal{X} \to [-\infty,+\infty]$ be a measurable function such that $\bar{\mu}(e^{\tau f}) \leq 1$. Then*

$$\sup_{t\geq 1} \frac{1}{t}\mathbf{E}\exp[t\,\pi_t(f)] < +\infty.$$

Proof. It is enough to prove the result in the case $\bar{\mu}(e^{\tau f}) = 1$. Then define $\bar{\mu}_f \in \mathcal{P}(\mathcal{X})$ as

$$\bar{\mu}_f(dx) := e^{\tau(x)\,f(x)}\bar{\mu}(dx).$$

Thus

$$\begin{aligned}\mathbf{E}\exp[t\,\pi_t(f)] &= \sum_{n=0}^{\infty}\mathbf{E}\exp\Big[\sum_{i=1}^{n}\tau(x_i)\,f(x_i) + (t-S_n)\,f(x_{n+1})\Big]\mathbb{1}_{\mathcal{N}_t=n} \\ &= \sum_{n=0}^{\infty}\int_{\mathcal{X}^{n+1}}\Big(\prod_{i=1}^{n}\bar{\mu}_f(dx_i)\Big)\bar{\mu}(dx_{n+1})\,\exp[(t-S_n)f(x_{n+1})]\mathbb{1}_{\mathcal{N}_t=n}.\end{aligned}$$

Note that $\{\mathcal{N}_t = n\} = \{S_n < t\}\cap\{\tau(x_{n+1}) \geq t - S_n\}$, so that denoting $\zeta_{n,f} \in \mathcal{P}([0,+\infty])$ the law of $S_n = \tau(x_1) + \ldots + \tau(x_n)$ with respect to $\prod_{i=1}^n \bar{\mu}_f(dx_i)$

$$\mathbf{E}\exp[t\,\pi_t(f)] = \sum_{n=0}^{\infty}\int_{[0,t[}\zeta_{n,f}(ds)\int_{\{\tau\geq t-s\}}\bar{\mu}(dx)e^{(t-s)f(x)}.$$

The rightest integral is bounded by 2, since $e^{(t-s)f(x)} \leq 1 + e^{\tau(x)f(x)}$ on $\{\tau \geq t-s\}$. Thus

$$\frac{1}{t}\mathbf{E}\exp[t\,\pi_t(f)] \leq \frac{2}{t}\sum_{n=0}^{\infty}\zeta_{n,f}([0,t)) = \frac{2}{t}\sum_{n=0}^{\infty}\mathbf{E}_f\mathbb{1}_{\mathcal{N}_t\geq n} = 2\,\mathbf{E}_f\frac{\mathcal{N}_t}{t},$$

where $\mathbf{E}_f$ denotes expectation with respect to $\bar{\mu}_f^{\otimes\mathbb{N}^+}$. By general renewal theory [1, Chapter V.4], $\mathbf{E}_f\mathcal{N}_t/t \to \frac{1}{\bar{\mu}_f(\tau)} < +\infty$ as $t \to +\infty$. □

Proof of Theorem 1.3, upper bound. Fix $\mathcal{O}$ an open subset of $\mathcal{P}(\mathcal{X})$. Then for each $f \in LSC(\mathcal{X})$ such that $\bar{\mu}(e^{\tau f}) \leq 1$

$$\begin{aligned}\frac{1}{t}\log\mathbf{P}_t(\mathcal{O}) &= \frac{1}{t}\log\mathbf{E}e^{-t\pi_t(f)}e^{t\pi_t(f)}\mathbb{1}_{\pi_t\in\mathcal{O}} \\ &\leq \frac{1}{t}\log\Big[e^{-t\inf_{\nu\in\mathcal{O}}\nu(f)}\mathbf{E}e^{t\pi_t(f)}\Big] = -\inf_{\nu\in\mathcal{O}}\nu(f) + \frac{1}{t}\log\mathbf{E}e^{t\pi_t(f)}.\end{aligned}$$

By taking the limsup $t \to \infty$, the last term in the above formula vanishes by Lemma 3.1. Optimizing over f

$$\varlimsup_t \frac{1}{t} \log \mathbf{P}_t(\mathcal{O}) \le -\sup\{\inf_{\nu \in \mathcal{O}} \nu(f), f \in LSC(\mathcal{X}) \,:\, \bar{\mu}(e^{\tau f}) \le 1\}. \tag{3.2}$$

Since (3.2) holds true for each open set $\mathcal{O} \subset \mathcal{X}$, and $\nu \mapsto \nu(f)$ is lower semicontinuous for $f \in LSC(\mathcal{X})$, the minimax lemma [4, Appendix 2, Lemma 3.3] yields

$$\varlimsup_t \frac{1}{t} \log \mathbf{P}_t(\mathcal{K}) \le -\inf_{\nu \in \mathcal{K}} \sup\{\nu(f), f \in LSC(\mathcal{X}) \,:\, \bar{\mu}(e^{\tau f}) \le 1\}$$

for each compact $\mathcal{K} \subset \mathcal{P}(\mathcal{X})$. By Lemma 2.4, the large deviations upper bound then holds true on compact sets. But closed sets are compact since $\mathcal{P}(\mathcal{X})$ is compact. $\square$

The following remark provides a standard approach for proving large deviations lower bounds, see for instance [8] and references therein.

Remark 3.2. If for each $\nu \in \mathcal{P}(\mathcal{X})$ there exists a sequence $(\mathbf{Q}_t)$ in $\mathcal{P}(\mathcal{P}(\mathcal{X}))$ such that $\lim_t \mathbf{Q}_t = \delta_\nu$ narrowly in $\mathcal{P}(\mathcal{P}(\mathcal{X}))$ and

$$\varlimsup_t \frac{1}{t} \mathbf{H}(\mathbf{Q}_t | \mathbf{P}_t) \le J(\nu),$$

then $(\mathbf{P}_t)_{t>0}$ satisfies a large deviations lower bound with rate given by the lower semicontinuous envelope of J.

For $t > 0$ let $\mathfrak{F}_t$ be the smallest σ-algebra on $\mathcal{X}^{\mathbb{N}^+}$ such that the map

$$\mathcal{X}^{\mathbb{N}^+} \ni \mathbf{x} \mapsto (x_1, \dots, x_{\mathcal{N}_t(\mathbf{x})+1}) \in \cup_{n \in \mathbb{N}^+} \mathcal{X}^n \hookrightarrow \mathcal{X}^{\mathbb{N}^+}$$

is Borel measurable. Note in particular that $\mathcal{N}_t \colon \mathcal{X}^{\mathbb{N}^+} \to \mathbb{N}$ and $\pi_t \colon \mathcal{X}^{\mathbb{N}^+} \to \mathcal{P}(\mathcal{X})$ are $\mathfrak{F}_t$ measurable (with respect to the discrete σ-algebra of $\mathbb{N}$ and the Borel σ-algebra on $\mathcal{P}(\mathcal{X})$ respectively).

Lemma 3.3. *Let $\mathcal{Y}$ be a Polish space, $F \colon \mathcal{X}^{\mathbb{N}^+} \to \mathcal{Y}$ a $\mathfrak{F}_t$-Borel measurable map, $(\bar{\mu}_i)_{i \in \mathbb{N}^+}$, $(\bar{\nu}_i)_{i \in \mathbb{N}^+}$ be sequences in $\mathcal{P}(\mathcal{X})$ and set $\Omega^\mu := \prod_{i \in \mathbb{N}^+} \bar{\mu}_i$, $\Omega^\nu := \prod_{i \in \mathbb{N}^+} \bar{\nu}_i$. Let $\mathbf{P}^F$, $\mathbf{Q}^F \in \mathcal{P}(\mathcal{Y})$ be the laws of F under Ω^μ and Ω^ν respectively. Then*

$$\mathbf{H}(\mathbf{Q}^F | \mathbf{P}^F) \le \sum_{j=1}^{\infty} \mathbf{H}(\bar{\nu}_j | \bar{\mu}_j)\, \Omega^\nu(\mathcal{N}_t \ge j-1).$$

In particular, if $\bar{\mu}_i = \bar{\mu}$ and $\bar{\nu}_i = \bar{\nu}$, then

$$\mathbf{H}(\mathbf{Q}^F | \mathbf{P}^F) \le \mathbf{H}(\bar{\nu} | \bar{\mu})\, \Omega^\nu(\mathcal{N}_t + 1).$$

Proof. For $r > 0$ let (as above) $h(r) = r(\log r - 1) + 1$, and let $\mathfrak{F}^F \subset \mathfrak{F}_t$ be the σ-algebra generated by F. Then for Ω^μ-a.e. $\mathbf{x}$

$$\frac{d\mathbf{Q}^F}{d\mathbf{P}^F}(F(\mathbf{x})) = \frac{d\Omega^\nu \circ F^{-1}}{d\Omega^\mu \circ F^{-1}}(F(\mathbf{x})) = \Omega^\mu\big(\frac{d\Omega^\nu}{d\Omega^\mu}\big|\mathfrak{F}^F\big)(\mathbf{x}).$$

Therefore changing variables in the integration and using the convexity of h

$$\begin{aligned} \mathbf{H}(\mathbf{Q}^F | \mathbf{P}^F) &= \int_{\mathcal{Y}} \mathbf{P}^F(dy)\, h\big(\frac{d\mathbf{Q}^F}{d\mathbf{P}^F}(y)\big) \\ &= \int_{\mathcal{X}^{\mathbb{N}^+}} \Omega^\mu(d\mathbf{x})\, h\big(\Omega^\mu\big(\frac{d\Omega^\nu}{d\Omega^\mu}\big|\mathfrak{F}^F\big)(\mathbf{x})\big) \le \int_{\mathcal{X}^{\mathbb{N}^+}} \Omega^\mu(d\mathbf{x})\, h\big(\Omega^\mu\big(\frac{d\Omega^\nu}{d\Omega^\mu}\big|\mathfrak{F}_t\big)(\mathbf{x})\big). \end{aligned}$$

For $n \in \mathbb{N}$, and $\mathbf{x}$ such that $\mathcal{N}_t(\mathbf{x}) = n$ one has $\Omega^\mu\big(\frac{d\Omega^\nu}{d\Omega^\mu}\big|\mathfrak{F}_t\big)(\mathbf{x}) = \prod_{j=1}^{n+1} \frac{d\nu_j}{d\mu_j}(x_j)$ and thus

$$\begin{aligned}
\mathbf{H}(\mathbf{Q}^F|\mathbf{P}^F) &\le \sum_{n\in\mathbb{N}} \int_{\mathcal{X}^{n+1}} \prod_{i=1}^{n+1} \mu_i(dx_i)\, h\Big(\prod_{j=1}^{n+1} \frac{d\nu_j}{d\mu_j}(x_j)\Big)\, \mathbb{1}_{\mathcal{N}_t(\mathbf{x})=n} \\
&= \sum_{n\in\mathbb{N}} \int_{\mathcal{X}^{n+1}} \prod_{i=1}^{n+1} \nu_i(dx_i)\, \log\Big(\prod_{j=1}^{n+1} \frac{d\nu_i}{d\mu_j}(x_j)\Big)\, \mathbb{1}_{\mathcal{N}_t(\mathbf{x})=n} \\
&= \sum_{j\in\mathbb{N}^+} \int_{\mathcal{X}^j} \prod_{i=1}^{j} \nu_i(dx_i) \log \frac{d\nu_j}{d\mu_j}(x_j)\, \mathbb{1}_{\mathcal{N}_t(\mathbf{x})\ge j-1}.
\end{aligned}$$

The event $\{\mathcal{N}_t(\mathbf{x}) \ge j-1\}$ only depends on $(x_1,\dots,x_{j-1})$. Therefore the last integral in the above formula splits into a product as

$$\mathbf{H}(\mathbf{Q}^F|\mathbf{P}^F) \le \sum_{j\in\mathbb{N}^+} \int_{\mathcal{X}^{j-1}} \prod_{i=1}^{j-1} \nu_i(dx_i)\, \mathbb{1}_{\mathcal{N}_t(\mathbf{x})\ge j-1} \int_{\mathcal{X}} \nu_j(dx_j)\, \log \frac{d\nu_j}{d\mu_j}(x_j)$$

which is easily rewritten as in the statement. □

Proof of Theorem 1.3, lower bound. In view of Proposition 2.6, and Remark 3.2, for each $\nu \in \mathcal{P}(\mathcal{X})$ such that $J(\nu) < +\infty$, one needs to find a sequence $(\mathbf{Q}_t)$ in $\mathcal{P}(\mathcal{P}(\mathcal{X}))$ such that $\mathbf{Q}_t \to \delta_\nu$ narrowly and $\overline{\lim}_t \frac{1}{t}\mathbf{H}(\mathbf{Q}_t|\mathbf{P}_t) \le J(\nu)$.

Fix a $\nu \in \mathcal{P}(\mathcal{X})$ absolutely continuous with respect to $\bar\mu$ and such that $\nu(1/\tau) \in\,]0,+\infty[$, and let $\Omega^\nu(d\mathbf{x}) := \prod_{i\in\mathbb{N}^+} \bar\nu(dx_i)$ as in Lemma 3.3. Set $\mathbf{Q}_t := \Omega^\nu \circ \pi_t^{-1}$. Since $\nu(1/\tau) < +\infty$, ergodic theorem yields $\lim_t \mathbf{Q}_t = \delta_\nu$. On the other hand, since π_t is $\mathfrak{F}_t$ measurable, one may apply Lemma 3.3 with $F = \pi_t$ to get

$$\frac{1}{t}\mathbf{H}(\mathbf{Q}_t|\mathbf{P}_t) \le \mathbf{H}(\bar\nu|\bar\mu)\frac{\Omega^\nu(\mathcal{N}_t+1)}{t}. \tag{3.3}$$

The renewal theorem [1, Chapter V.4] implies $\lim_t \Omega^\nu(\mathcal{N}_t)/t = \nu(1/\tau)$, which concludes the proof. □

References

[1] S. Asmussen, *Applied Probability and Queues*, Second Edition, Application of Mathematics **51**, Springer–Verlag, New York (2003). MR-1978607

[2] A. Dembo, O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlett Publishers (1993). MR-1202429

[3] Donsker M. D., Varadhan, S. R. S., *Asymptotic evaluation of certain Markov process expectations for large time. I. II.* Comm. Pure Appl. Math. **28** (1975), 1–47; ibid. **28**, 279–301 (1975). MR-0386024

[4] Kipnis C., Landim C., Scaling limits of interacting particle systems. Springer-Verlag, Berlin (1999). MR-1707314

[5] R. Lefevere, M. Mariani, L. Zambotti, (2011) *Large deviations for renewal processes*, Stochastic Processes and Their Applications, Volume 121, Issue 10, 2243-2271. MR-2822776

[6] R. Lefevere, M. Mariani and L. Zambotti, *Large deviations of the current in stochastic collisional dynamics*, J. Math. Phys. 52 (2011). MR-2814865

[7] R. Lefevere, M. Mariani, L. Zambotti, *Large deviations for a random speed particle*, ALEA, Vol. IX, pages 739-760 (2012). MR-3069383

[8] M. Mariani , A Γ-convergence approach to large deviations Ann. Sc. Norm. Super. Pisa Cl. Sci. (to appear).

[9] D. Stroock, Probability theory. An analytic view, 2nd edition. Cambridge University Press, (2011). MR-2760872

[10] B. Tsirelson, *From uniform renewal theorem to uniform large and moderate deviations for renewal-reward processes*, Electronic Communications in Probability 18:52, 1-13 (2013). MR-3078015